





CC -!- SIMILARITY: TO 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONGLICININ, ETC.).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L34402; PAB00861.1; -.

DR HSSP; P50477; ICAX.

DR INTERPRO; IPR001113; -.

DR PFAM; PF00546; Seedstore\_7s; 1.

KW Allergen.

CC SEQUENCE 626 AA; 71345 MW; 1A6BBBE41490D0E3 CRC64; SQ

Query Match 21.9%; Score 69; DB 1; Length 246; Best Local Similarity 32.1%; Pred: No. 0.99; Mismatches 17; Indels 8; Gaps 2; Matches 17; Conservative 11; Mismatches 17; Indels 8; Gaps 2;

QY 2 NPKHNCIQLQSCNSRSRDSYRNQACHARNLKVKEECEGEIP--RRRRPQH 52

Db 100 NPKKPKEEKRKERHHKHKHRH---KEKRELOEQEKPDKPKERPKH 146

RESULT 6

VC1B\_GOSHI 6

ID VC1B\_GOSHI STANDARD PRT; 588 AA.

AC P09801;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

RT 15-JUL-1999 (Rel. 38, Last annotation update)

DE VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B).

OS Gossypium hirsutum (upland cotton).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; Malvales; Malvaceae; Gossypium.

OC Malvales; Malvaceae; Gossypium.

RN [1]

RP SEQUENCE FROM N.A.

RA Chian C.A., Pyle J.B., Legocki A.B., Dure L. III;

RT "developmental biochemistry of cottonseed embryogenesis and germination. XVIII. cDNA and amino acid sequences of the members of the storage protein families."

RL Plant Mol. Biol. 7:475-489(1986);

CC -!- FUNCTION: SEED STORAGE PROTEIN.

CC -!- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.

CC -!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONGLICININ, CONGLICININ, ETC.).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M16891; AA033071.1; -.

DR PIR; A30838; FWCNAB.

DR HSSP; P50477; ICAX.

DR INTERPRO; IPR001113; -.

DR PFAM; PF00546; Seedstore\_7s; 1.

KW Seed storage protein; Signal.

FT SIGNAL 1

FT CHAIN 25

FT SEQUENCE 588 AA; 69729 MW; 63E699B29AB8ADEB CRC64;

CC -----

CC -!- SIMILARITY: BELONGS TO THE TONB FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U62555; RAB09550.1; -.

KW Transport; Protein transport; Inner membrane; Periplasmic;

FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 8 28 ANCHOR (POTENTIAL).

FT DOMAIN 29 246 PERIPLASMIC (POTENTIAL).

FT SEQUENCE 246 AA; 27785 MW; C9532F619CBAB55 CR64;

CC -----

CC -!- SIMILARITY: CONGLICININ, ETC.).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L34402; PAB00861.1; -.

DR HSSP; P50477; ICAX.

DR INTERPRO; IPR001113; -.

DR PFAM; PF00546; Seedstore\_7s; 1.

KW Allergen.

CC SEQUENCE 626 AA; 71345 MW; 1A6BBBE41490D0E3 CRC64; SQ

Query Match 21.9%; Score 69; DB 1; Length 246; Best Local Similarity 32.1%; Pred: No. 0.99; Mismatches 19; Indels 17; Gaps 5; Matches 19; Conservative 11; Mismatches 19; Indels 17; Gaps 5;

QY 1 ENP--KHNKCIQLQSCNSRSRDSYRNQACHARNLKVKEECEGEIPPRPR 49

Db 81 EDPQRYYEECQECRQEEERORPQ-CQORC-LKRFEGEQDOSORQFECQO---HCHQ 134

QY 50 P0HPER 55

Db 135 E0RPER 140

DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HYPOTHETICAL PROTEIN T20G5.3 IN CHROMOSOME III (FRAGMENT).  
 GN T20G5..3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peioderinae; Caenorhabditis.  
 RL [1]  
 RN SEQUENCE FROM N. A.  
 RP STRAIN-BRISTOL N2;  
 RA Berks M., Smith A.;  
 CC Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.  
 -!- SIMILARITY: CONTAINS 45 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL: Z30423; CAM83007.1; -.  
 DR PIR: S42373; S42373.  
 DR WORMPEP; P35555; IEMO.  
 DR WORMPEP; T20G5..3; CE00478.  
 DR INTERPRO; IPR00082; -.  
 DR INTERPRO; IPR00152; -.  
 DR INTERPRO; IPR00561; -.  
 DR INTERPRO; IPR01881; -.  
 DR INTERPRO; IPR02035; -.  
 DR PROSITE; PS01186; EGF-1; 5.  
 DR PROSITE; PS00008; EGF-31.  
 DR PROSITE; PF01390; SEA; 2.  
 DR PROSITE; PF00052; vwa; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 28.  
 DR PROSITE; PS00022; EGF-1; 1.  
 DR PROSITE; PS01187; EGF-CA; 1.  
 KW Hypothetical protein; EGF-like domain; Repeat; Transmembrane.  
 FT NON\_TER 1  
 FT DOMAIN <1 2701 EXTRACELLULAR (POTENTIAL).  
 FT TRANSEM 2702 2722 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 2723 3051 EGF-LIKE 1 (PARTIAL).  
 FT DOMAIN <1 44 EGF-LIKE 2.  
 FT DOMAIN 46 94 EGF-LIKE 3.  
 FT DOMAIN 100 144 EGF-LIKE 4.  
 FT DOMAIN 145 192 EGF-LIKE 5.  
 FT DOMAIN 194 245 EGF-LIKE 6.  
 FT DOMAIN 247 296 EGF-LIKE 7.  
 FT DOMAIN 313 354 EGF-LIKE 8.  
 FT DOMAIN 355 402 EGF-LIKE 9.  
 FT DOMAIN 404 452 EGF-LIKE 10.  
 FT DOMAIN 454 503 EGF-LIKE 10.  
 FT DOMAIN 514 699 VWA.  
 FT DOMAIN 705 750 EGF-LIKE 11.  
 FT DOMAIN 750 794 EGF-LIKE 12.  
 FT DOMAIN 805 846 EGF-LIKE 13.  
 FT DOMAIN 847 892 EGF-LIKE 14.  
 FT DOMAIN 892 940 EGF-LIKE 15.  
 FT DOMAIN 942 990 EGF-LIKE 16.  
 FT DOMAIN 992 1039 EGF-LIKE 17.  
 FT DOMAIN 1043 1091 EGF-LIKE 18.  
 FT DOMAIN 1093 1144 EGF-LIKE 19.  
 FT DOMAIN 1146 1195 EGF-LIKE 20.  
 FT DOMAIN 1197 1247 EGF-LIKE 21.  
 FT DOMAIN 1247 1295 EGF-LIKE 22.  
 FT DOMAIN 1298 1346 EGF-LIKE 23.  
 FT DOMAIN 1352 1396 EGF-LIKE 24.  
 FT DOMAIN 1397 1444 EGF-LIKE 25.  
 FT DOMAIN 1446 1492 EGF-LIKE 26.  
 FT DOMAIN 1494 1542 EGF-LIKE 27.  
 FT DOMAIN 1544 1592 EGF-LIKE 28.  
 FT DOMAIN 1594 1642 EGF-LIKE 29.  
 FT DOMAIN 1644 1692 EGF-LIKE 30.  
 FT DOMAIN 1693 1739 EGF-LIKE 31.  
 FT DOMAIN 1740 1788 EGF-LIKE 32.  
 FT DOMAIN 1797 1847 EGF-LIKE 33.  
 FT DOMAIN 1849 1900 EGF-LIKE 34.  
 FT DOMAIN 1902 1950 EGF-LIKE 35.  
 FT DOMAIN 1952 1998 EGF-LIKE 36.  
 FT DOMAIN 2000 2047 EGF-LIKE 37.  
 FT DOMAIN 2047 2095 EGF-LIKE 38.  
 FT DOMAIN 2117 2156 EGF-LIKE 39.  
 FT DOMAIN 2155 2284 SEA 1.  
 FT DOMAIN 2232 2332 EGF-LIKE 40.  
 FT DOMAIN 2231 2457 SEA 2.  
 FT DOMAIN 2460 2504 EGF-LIKE 41.  
 FT DOMAIN 2508 2556 EGF-LIKE 42.  
 FT DOMAIN 2556 2608 EGF-LIKE 43.  
 FT DOMAIN 2657 2693 EGF-LIKE 44, CALCIUM-BINDING (POTENTIAL).  
 FT DISULFID 2 13 BY SIMILARITY.  
 FT DISULFID 7 22 BY SIMILARITY.  
 FT DISULFID 24 43 BY SIMILARITY.  
 FT DISULFID 50 63 BY SIMILARITY.  
 FT DISULFID 57 72 BY SIMILARITY.  
 FT DISULFID 74 93 BY SIMILARITY.  
 FT DISULFID 104 120 BY SIMILARITY.  
 FT DISULFID 112 129 BY SIMILARITY.  
 FT DISULFID 131 143 BY SIMILARITY.  
 FT DISULFID 149 163 BY SIMILARITY.  
 FT DISULFID 157 172 BY SIMILARITY.  
 FT DISULFID 174 191 BY SIMILARITY.  
 FT DISULFID 198 214 BY SIMILARITY.  
 FT DISULFID 208 223 BY SIMILARITY.  
 FT DISULFID 225 244 BY SIMILARITY.  
 FT DISULFID 251 265 BY SIMILARITY.  
 FT DISULFID 259 274 BY SIMILARITY.  
 FT DISULFID 276 295 BY SIMILARITY.  
 FT DISULFID 317 330 BY SIMILARITY.  
 FT DISULFID 324 339 BY SIMILARITY.  
 FT DISULFID 341 353 BY SIMILARITY.  
 FT DISULFID 359 371 BY SIMILARITY.  
 FT DISULFID 365 380 BY SIMILARITY.  
 FT DISULFID 382 401 BY SIMILARITY.  
 FT DISULFID 408 421 BY SIMILARITY.  
 FT DISULFID 415 430 BY SIMILARITY.  
 FT DISULFID 432 451 BY SIMILARITY.  
 FT DISULFID 458 472 BY SIMILARITY.  
 FT DISULFID 466 481 BY SIMILARITY.  
 FT DISULFID 483 502 BY SIMILARITY.  
 FT DISULFID 709 725 BY SIMILARITY.  
 FT DISULFID 717 734 BY SIMILARITY.  
 FT DISULFID 736 749 BY SIMILARITY.  
 FT DISULFID 754 768 BY SIMILARITY.  
 FT DISULFID 762 778 BY SIMILARITY.  
 FT DISULFID 780 793 BY SIMILARITY.  
 FT DISULFID 809 822 BY SIMILARITY.  
 FT DISULFID 816 831 BY SIMILARITY.  
 FT DISULFID 833 845 BY SIMILARITY.  
 FT DISULFID 851 867 BY SIMILARITY.  
 FT DISULFID 859 876 BY SIMILARITY.  
 FT DISULFID 878 891 BY SIMILARITY.  
 FT DISULFID 896 909 BY SIMILARITY.  
 FT DISULFID 903 918 BY SIMILARITY.  
 FT DISULFID 920 939 BY SIMILARITY.  
 FT DISULFID 946 959 BY SIMILARITY.  
 FT DISULFID 953 968 BY SIMILARITY.  
 FT DISULFID 970 989 BY SIMILARITY.  
 FT DISULFID 996 1010 BY SIMILARITY.  
 FT DISULFID 1004 1019 BY SIMILARITY.  
 FT DISULFID 1021 1038 BY SIMILARITY.  
 FT DISULFID 1047 1060 BY SIMILARITY.  
 FT DISULFID 1054 1070 BY SIMILARITY.  
 FT DISULFID 1072 1090 BY SIMILARITY.  
 FT DISULFID 1097 1113 BY SIMILARITY.





CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- ALTERNATIVE PRODUCTS: MULTIPLE ISOFORMS ARE CREATED BY ALTERNATIVE  
 CC SPLICING EVENTS, WHICH SEEM TO OCCUR IN A TISSUE-SPECIFIC MANNER.  
 CC THE SEQUENCE SHOWN HERE IS THAT OF CACM4A. THE REGION SEQUENCED IN  
 CC ROB3 AND RKC5 IS IDENTICAL TO CACM4.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PANCREATIC ISLETS AND B-  
 CC LYMPHOCTYES.  
 CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE  
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE  
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS  
 CC PROBABLY REPRESENT THE VOLTAGE SENSOR AND ARE CHARACTERIZED BY A  
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.  
 CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS  
 CC FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC  
 DR EMBL: D38101; BAA07282.1; -.  
 DR EMBL: D38102; BAA07283.1; -.  
 DR EMBL: M57682; AA42015.1; -.  
 DR EMBL: U44005; AA460515.1; -.  
 DR EMBL: M92221; AA40895.1; -.  
 DR EMBL: U31772; AA89156.1; -.  
 DR EMBL: U49126; AAB01634.1; -.  
 DR EMBL: U49127; AAB01635.1; -.  
 DR EMBL: U49128; AAB01636.1; -.  
 DR EMBL: U49129; AAB01637.1; -.  
 DR EMBL: U49130; AAB01638.1; -.  
 DR EMBL: U49131; AAB01639.1; -.  
 DR EMBL: U49132; AAB01640.1; -.  
 DR EMBL: U49133; AAB01641.1; -.  
 DR EMBL: U49134; AAB01642.1; -.  
 DR EMBL: U49135; AAB01643.1; -.  
 DR EMBL: U49136; AAB01644.1; -.  
 DR EMBL: U49137; AAB01645.1; -.  
 DR EMBL: U49138; AAB01646.1; -.  
 DR EMBL: U49139; AAB01647.1; -.  
 DR EMBL: U49140; AAB01648.1; -.  
 DR EMBL: U49141; AAB01649.1; -.  
 DR EMBL: U49142; AAB01650.1; -.  
 DR EMBL: U49143; AAB01651.1; -.  
 DR EMBL: U49144; AAB01652.1; -.  
 DR EMBL: U49145; AAB01653.1; -.  
 DR EMBL: U49146; AAB01654.1; -.  
 DR EMBL: U49147; AAB01655.1; -.  
 DR EMBL: U49148; AAB01656.1; -.  
 DR EMBL: U49149; AAB01657.1; -.  
 DR EMBL: U49150; AAB01658.1; -.  
 DR EMBL: U49151; AAB01659.1; -.  
 DR EMBL: U49152; AAB01660.1; -.  
 DR EMBL: U49153; AAB01661.1; -.  
 DR EMBL: U49154; AAB01662.1; -.  
 DR EMBL: U49155; AAB01663.1; -.  
 DR EMBL: U49156; AAB01664.1; -.  
 DR EMBL: U49157; AAB01665.1; -.  
 DR EMBL: U49158; AAB01666.1; -.  
 DR EMBL: U49159; AAB01667.1; -.  
 DR EMBL: U49160; AAB01668.1; -.  
 DR EMBL: U49161; AAB01669.1; -.  
 DR EMBL: U49162; AAB01670.1; -.  
 DR EMBL: U49163; AAB01671.1; -.  
 DR EMBL: U49164; AAB01672.1; -.  
 DR EMBL: U49165; AAB01673.1; -.  
 DR EMBL: U49166; AAB01674.1; -.  
 DR EMBL: U49167; AAB01675.1; -.  
 DR EMBL: U49168; AAB01676.1; -.  
 DR EMBL: U49169; AAB01677.1; -.  
 DR EMBL: U49170; AAB01678.1; -.  
 DR EMBL: U49171; AAB01679.1; -.  
 DR EMBL: U49172; AAB01680.1; -.  
 DR EMBL: U49173; AAB01681.1; -.  
 DR EMBL: U49174; AAB01682.1; -.  
 DR EMBL: U49175; AAB01683.1; -.  
 DR EMBL: U49176; AAB01684.1; -.  
 DR EMBL: U49177; AAB01685.1; -.  
 DR EMBL: U49178; AAB01686.1; -.  
 DR EMBL: U49179; AAB01687.1; -.  
 DR EMBL: U49180; AAB01688.1; -.  
 DR EMBL: U49181; AAB01689.1; -.  
 DR EMBL: U49182; AAB01690.1; -.  
 DR EMBL: U49183; AAB01691.1; -.  
 DR EMBL: U49184; AAB01692.1; -.  
 DR EMBL: U49185; AAB01693.1; -.  
 DR EMBL: U49186; AAB01694.1; -.  
 DR EMBL: U49187; AAB01695.1; -.  
 DR EMBL: U49188; AAB01696.1; -.  
 DR EMBL: U49189; AAB01697.1; -.  
 DR EMBL: U49190; AAB01698.1; -.  
 DR EMBL: U49191; AAB01699.1; -.  
 DR EMBL: U49192; AAB01700.1; -.  
 DR EMBL: U49193; AAB01701.1; -.  
 DR EMBL: U49194; AAB01702.1; -.  
 DR EMBL: U49195; AAB01703.1; -.  
 DR EMBL: U49196; AAB01704.1; -.  
 DR EMBL: U49197; AAB01705.1; -.  
 DR EMBL: U49198; AAB01706.1; -.  
 DR EMBL: U49199; AAB01707.1; -.  
 DR EMBL: U49200; AAB01708.1; -.  
 DR PRINTS: PRO0167; CACHANNEL.  
 KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;  
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;  
 KW Calcium-binding; Phosphorylation; Alternative splicing.  
 FT REPEAT 528 774 II.  
 FT REPEAT 892 1174 III.  
 FT REPEAT 1211 1486 IV.  
 FT DOMAIN 127 145 S1 OF REPEAT I (POTENTIAL).  
 FT DOMAIN 146 163 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 164 183 S2 OF REPEAT I (POTENTIAL).  
 FT DOMAIN 184 195 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 196 214 S3 OF REPEAT I (POTENTIAL).  
 FT DOMAIN 215 235 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 236 254 S4 OF REPEAT I (POTENTIAL).  
 FT DOMAIN 255 273 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 274 293 S5 OF REPEAT I (POTENTIAL).  
 FT DOMAIN 294 381 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 382 406 S6 OF REPEAT I (POTENTIAL).  
 FT DOMAIN 407 582 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 583 602 S1 OF REPEAT II (POTENTIAL).  
 FT DOMAIN 603 617 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 618 636 S2 OF REPEAT II (POTENTIAL).  
 FT DOMAIN 637 644 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 645 663 S3 OF REPEAT II (POTENTIAL).  
 FT DOMAIN 664 673 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 674 692 S4 OF REPEAT II (POTENTIAL).  
 FT DOMAIN 693 711 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 712 732 S5 OF REPEAT II (POTENTIAL).  
 FT DOMAIN 733 787 811 845 876 912 945 964 980 1000 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 787 811 845 876 912 945 964 980 1000 S6 OF REPEAT II (POTENTIAL).  
 FT TRANSMEM 811 845 876 912 945 964 980 1000 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 845 876 912 945 964 980 1000 S1 OF REPEAT III (POTENTIAL).  
 FT TRANSMEM 876 912 945 964 980 1000 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 912 945 964 980 1000 S2 OF REPEAT III (POTENTIAL).  
 FT TRANSMEM 945 964 980 1000 S3 OF REPEAT III (POTENTIAL).  
 FT TRANSMEM 964 980 1000 S4 OF REPEAT III (POTENTIAL).  
 FT TRANSMEM 980 1000 S5 OF REPEAT III (POTENTIAL).  
 FT TRANSMEM 1000 S6 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1001 1012 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1013 1031 S3 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1032 1057 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1038 1057 S4 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1058 1076 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1077 1096 S5 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1097 1186 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 11207 1207 S6 OF REPEAT III (POTENTIAL).  
 FT DOMAIN 1208 1264 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1265 1283 S1 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1284 1298 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1299 1318 S2 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1319 1325 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1326 1347 S3 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1348 1357 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1358 1377 S4 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1378 1396 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1397 1416 S5 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1417 1483 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1484 1508 S6 OF REPEAT IV (POTENTIAL).  
 FT DOMAIN 1509 2203 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1 7 POLY-MET.  
 FT DOMAIN 712 718 POLY-LEU.  
 FT DOMAIN 886 897 BINDING TO THE BETA SUBUNIT (BY SIMILARITY).  
 FT DOMAIN 429 446 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).  
 FT SITE 364 364 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).  
 FT SITE 763 763 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).  
 FT SITE 1160 1160 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).  
 FT SITE 1450 1450 CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).  
 FT BINDING 1134 1224 TO DIHYDROPYRIDINES (BY SIMILARITY).  
 FT BINDING 1464 1530 TO DIHYDROPYRIDINES (BY SIMILARITY).

Query Match 19.0%; Score 60; DB 1; Length 2203;  
 Best Local Similarity 34.9%; Prede. No. 65; 9; Mismatches 18; Indels 14; Gaps 4;  
 Matches 22; Conservative 9;  
 QY 1 ENPKHKN--CTQSCNSR---DSYRNQA---CHARCNLKVKEBCEBIPR---P 46  
 DB 843 ENPKKNNKPEVQJANSNDNKVITYDQEEAEDKDPYPPCDVPGEEEEEDEPEVPGP 902  
 QY 47 RPR 49  
 DE 903 RPR 905  
 Db 903 RPR 905  
 RESULT 11  
 ID BARI\_CHITE STANDARD: PRT; 174 AA.  
 ID BARI\_CHITE AC  
 ID BARI\_CHITE P02849;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE BALBIANI RING PROTEIN 1 (Giant Secretory Protein 1-A) (GSP-1A)  
 DE (FRAGMENT).  
 GN BRI.  
 OS Chironomus tentans (Midge).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neopera; Endopterygota; Diptera; Hematocera;  
 OC Chironomoidea; Chironomidae; Chironominae; Chironomus.  
 RN [1] SEQUENCE OF 1-116 FROM N. A. (CLONE PCTBR1-1).  
 RC TISSUE-SALIVARY GLAND;  
 RX MEDLINE=8338361; PubMed=6305953;  
 RA Case S.T., Byers M.R.;  
 RT "Repeated nucleotide sequence arrays in Balbiani ring 1 of Chironomus  
 RT tentans contain internally nonrepeating and subrepeating elements.";  
 RL J. Biol. Chem. 258:7793-7799 (1983).  
 [2]

RA TISSUE-SALIVARY GLAND; FROM N.A. (CLONE PCT21).

RC Wieslander L., Sumegi J., Daneholt B.;

RT "Evidence for a common ancestor sequence for the Balbiani ring 1 and

RT Balbiani ring 2 genes in *Chironomus tentans*."

RL Proc. Natl. Acad. Sci. U.S.A. 79:6956-6960(1982).

CC -I- STRUCTURE: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR

CC -I- SUBCELLULAR LOCATION: THE LARVAL TUBE

CC -I- SUBCELLULAR LOCATION: SECRETED

CC -I- TISSUE SPECIFICITY: SALIVARY GLAND.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: K00447; AAH28238.1; -

DR EMBL; J01055; AAH28236.1; -

KW PIR; A03339; BQC1T.

FT repeat.

FT NON-TER 1 1

FT DOMAIN 42 85 4 X 11 AA TANDEM REPEATS.

FT REPEAT 42 52 1-1.

FT REPEAT 53 63 1-2.

FT REPEAT 64 74 1-3.

FT REPEAT 75 85 1-4.

FT DOMAIN 124 167 4 X 11 AA TANDEM REPEATS.

FT REPEAT 124 134 2-1.

FT REPEAT 135 145 2-2.

FT REPEAT 146 156 2-3.

FT REPEAT 157 167 2-4.

FT VARIANT 40 40 N -> K (IN CLONE PCT21).

FT VARIANT 57 57 G -> E (IN CLONE PCT21).

FT VARIANT 72 72 K -> R (IN CLONE PCT21).

FT VARIANT 86 86 R -> G (IN CLONE PCT21).

FT NON-TER 174 174 AA; 18920 MW; E86E16C6413D751 CRC64;

SQ

Query Match 18.9%; Score 59.5; DB 1; Length 174;

Best Local Similarity 26.5%; Pred. No. 7; 4;

Matches 18; Conservative 11; Mismatches 24; Indels 15; Gaps 4;

Qy 3 PKHNC--QSCNSERDSYRN--OACHARNLL--KVEKECEERGEIIPR-----R 47

Db 82 PRPERGSGAMRKAEAKCARRNGRNASKCCTSGAKPSKSEPSKGSKPRPEKSKESK 141

Qy 48 PRQHPER 55

Db 142 PRPEKPSK 149

RESULT 12

MSTL\_DROHY STANDARD; PRT; 344 AA.

AC 008695; 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE AXNONE-ASSOCIATED PROTEIN MST101(1).

DE GN MST101(1).

OS Drosophila hydei (Fruit fly).

OC Drosophila; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Eukaryota; Metazoa; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydriidae; Drosophilidae; Drosophila.

OC [1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

CC TISSUE=TESTIS;

CC MEDLINE=94200512; PubMed=8150205;

CC Neesen J., Bueltemann H., Heinlein U.A.;

RT "The drosophila hydei gene Dmms101(1) encodes a testis-specific, repetitive, axoneme-associated protein with differential abundance in Y chromosomal deletion mutant flies."

RT Dev. Biol. 162:414-425(1994).

RL ASSOCIATED WITH AXONEMAL STRUCTURES.

CC -I- SUBCELLULAR LOCATION: TISSUE SPECIFICITY: TESTIS. LOCATED IN SPERMATOCTES AND SPERMATID BUNDLES.

CC -I- DOMAIN: THE PREDOMINANT STRUCTURE IS ALPHA-HELICAL.

CC -I- POLYMORPHISM: THE NUMBER OF REPEATS VARIES BETWEEN STRAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X73480; CA551875.1; -

DR PIR; S34153; S34153.

DR HSSP; P01032; IC5A.

DR FLYBASE; FBgn0011816; Dhyd\mst101(1).

FT DOMAIN 58 337 19 X 16 AA APPROXIMATE TANDEM REPEATS OF K-K-C-X-A-[KQ]-K-X-E-X-A-X.

KW Sperm; Repeat; Multienzyme family; Polymorphism.

FT SEQUENCE 344 AA; 37793 MW; 24C65D2510387E2A CRC64;

SQ

Query Match 18.7%; Score 59; DB 1; Length 344;

Best Local Similarity 28.6%; Pred. No. 15;

Matches 12; Conservative 9; Mismatches 19; Indels 2; Gaps 1;

Qy 1 ENPKHNCQSCNSERDSYRNQACHARNLL--KVEKECEERGEIIPR-----R 40

Db 191 EAEEKKCAEAKKKEAERKKCEBRAKKKEAAKKKCCE 232

RESULT 13

TSR4\_GIALA STANDARD; PRT; 713 AA.

AC P21849; 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 15-JUL-1999 (Rel. 33, Last annotation update)

DE MAJOR SURFACE-LABELED TROPHOZOITE ANTIGEN PRECURSOR.

GN TSA 417.

OS Giardia lamblia (Giardia intestinalis).

OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.

RN [1]

RP SEQUENCE FROM N.A. / WB;

RC STRAIN=ATCC 30957 / WB;

RX MEDLINE=90280395; PubMed=2352929;

RX MEDLINE=90280395; PubMed=2352929;

RA Gillin F.D., Hagblom P., Harwood J., Aley S.B., Reiner D.S., McCaffery M., So M., Guiney D.G.;

RA RT "Isolation and expression of the gene for a major surface protein of

RT Giardia lamblia";

RT Prog. Natl. Acad. Sci. U.S.A. 87:4463-4467(1990).

RN [2]

RP SEQUENCE OF 480-620 FROM N.A.

RC STRAIN=AD-1;

RX MEDLINE=93314970; PubMed=8325510;

RA EY P.L., Mayrhofer G.;

RT "Two genes encoding homologous 70-kDa surface proteins are present

RT within individual trophozoites of the binucleate protozoan parasite

RT Giardia intestinalis";

RT Gene 129:257-262(1993).

RL [1]

CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -I- DOMAIN: CONTAINS 29 REPEATS OF THE CX(C) MOTIF.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC purposes requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [bioinf@ebi.ac.uk](mailto:bioinf@ebi.ac.uk)

DR	PRINTS;	PR00237;	GPCRRHODOPSN.
DR	PROSITE;	PS00237;	G-PROTEIN_RECIP_FL1_1.
DR	PROSITE;	PS05262;	G-PROTEIN_RECIP_FL1_2.
KW	G-PROTEIN coupled receptor;	Transmembrane;	Glycoprotein.
FT	DOMAIN	1	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	29	POTENTIAL.
FT	DOMAIN	54	CYTOSPLASMIC (POTENTIAL).
FT	TRANSMEM	65	POTENTIAL.
FT	DOMAIN	88	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	103	POTENTIAL.
FT	DOMAIN	125	CYTOSPLASMIC (POTENTIAL).
FT	TRANSMEM	148	POTENTIAL.
FT	DOMAIN	168	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	240	POTENTIAL.
FT	DOMAIN	260	CYTOSPLASMIC (POTENTIAL).
FT	TRANSMEM	521	POTENTIAL.
FT	DOMAIN	546	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	552	POTENTIAL.
FT	DOMAIN	576	CYTOSPLASMIC (POTENTIAL).
FT	DISULFID	101	BY SIMILARITY.
FT	CARBOHYD	178	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	207	N-LINKED (GICNAC. . .) (POTENTIAL).
FT	CARBOHYD	215	N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ	SEQUENCE	638 AA;	70000 MW;
			65FA928B5C01D34F CRC64;

Query Match	18.6%	Score	58.5%	DB	1	Length	638
Best Local Similarity	29.6%	Pred. No.	30				
Matches 16;	Conservative	8;	Mismatches	17;	Indels	13;	Gaps
							3

Search completed: March 1, 2001, 16:20:36  
Job time: 211 sec

Job time: 211 sec